

VARIATION OF BIVARIATE CHARACTERS FROM THE STANDPOINT OF ALLOMETRY

by *ITARU HAYAMI and AKIHIKO MATSUKUMA*

ABSTRACT. In fossil biometry simple ratios between two linear measurements have been frequently applied as a third variable for representation of the variation of a bivariate character. Theoretically, however, it is obvious that its frequency distribution is sometimes strongly influenced by the heterogeneity of sample, especially the age distribution of a fossil population, and is apt to be artificially skewed and platykurtic. In order to analyse the real state of frequency distribution of bivariate characters and to apply further advanced statistical techniques for taxonomic identification and discrimination, it may be necessary to use some other index which is little, or preferably not at all influenced by growth. In this respect the diagonal distance from each point to the reduced major axis on a double logarithmic scatter diagram may be a more desirable index, if a single power function represents the relative growth of the organism. The advantage of this method is also recognized empirically by a comparative study on an actual fossil sample of *Glycymeris rotunda* from the Pliocene of central Japan. Some comments are given as to the definition of isometry and allometry.

MODERN palaeontology focuses on the population rather than the individual. A biometrical study of individual variation seems to be important, primarily because it may offer fundamental and objective information for the consideration of classification and evolution. There are various kinds of individual variation which are controlled by different factors. As classified by Mayr, Linsley, and Usinger (1953) and again by Mayr (1969), some are genetic and the others are non-genetic. Here we intend to discuss mainly non-sex-associated continuous variation and methods for the representation of bivariate characters in a population. It is no doubt important to consider from a biological viewpoint whether the variation is genetic or non-genetic. Unfortunately, however, the distinction is usually not easy in fossils. We presume that such bivariate characters as are discussed here might be controlled both by genes and environments.

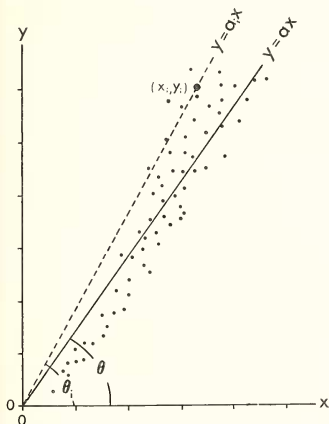
Various kinds of characters may be used for the representation of continuous variation. In neontology individual variation can be well recognized on the basis of a univariate character, if the population proves to be composed of individuals of almost the same age. This method is logically appropriate and certainly applicable also in palaeontology, if the character does not change as the organism grows. For instance, the number of simple radial ribs of bivalves can be regarded as a good character. Such characters are, however, usually restricted to count, whereas linear measurements themselves are hardly applicable as the growth-invariant index.

Many previous authors have shown histograms of bivariate characters in order to express the individual variation, taking the simple ratio between pairs of linear measurements as the index. Ratios are, in fact, often more useful than linear measurements in taxonomy. It is, however, obvious that this conventional method bears some theoretical difficulties, because such a ratio may change to some extent through the growth of organism. In this respect the dynamic concept of relative growth should be introduced in the study of individual variation. We have been devoting ourselves to seek good indices for the representation of bivariate characters from the standpoint of allometry.

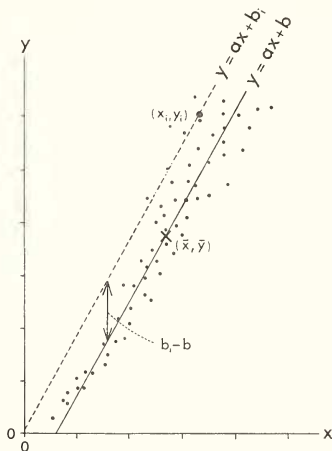
As such indices may deserve general application in palaeontology, we intend to discuss the problem in some detail and to evaluate them on the basis of some actual data.

PROBLEMS

When one intends to discuss the continuous variation of a sample and to apply any advanced statistical techniques for the estimation of the nature of a population, it may be necessary as an initial condition to recognize or presume normal frequency distribution of characters. Therefore, one should select the most adequate and meaningful index for each character especially carefully and strictly in order to represent the variation with the combination of two variables.



TEXT-FIG. 1. Hypothetical scatter diagram showing the concept of isometric variation, where the simple ratio y/x is applied as an index. $a_i - a$ or $\tan \theta_i - \tan \theta$ means the deviation for an arbitrary individual.



TEXT-FIG. 2. Hypothetical scatter diagram showing the concept of 'initial index variation'. The line $y = ax + b$ is the reduced major axis for a sample. $b_i - b$ means the deviation for an arbitrary individual.

The ratio between two linear measurements, which is the simplest index, has been used frequently in fossil biometry. Many authors distinguished one species from another on the basis of the difference of mean simple ratios. Some biometricians applied Student's t -test for the discrimination of populations. In this case, as shown in text-fig. 1, the relation of a linear equation:

$$y = a_i x \quad (1)$$

is presumed for the growth of each individual. If the average value of a_i is defined as a , the growth of a typical individual is represented by the following equation:

$$y = ax \quad (2)$$

In this case $a_i - a$ or $\tan \theta_i - \tan \theta$ is regarded as the deviation of bivariate character from the mean. We may provisionally call this expression *isometric variation*.

Some authors (e.g. Omori and Utashiro 1954) suggested another statistical method to show the variation of a sample, where the following linear equation was regressed by means of the least-squares criterion from all the given individuals:

$$y = ax + b. \quad (3)$$

They presumed that the slope a is nearly constant in one species and that the initial index (y -intercept) b is a variable mainly related to the intraspecific or geographic variation. In that method, although the presumption of the constancy of a is quite dubious, the growth of each individual would be expressed by the following formula:

$$y = ax + b_i \quad (4)$$

where b_i or $b_i - b$ is regarded as the index for the variation (text-fig. 2).

Such an estimation of variation on the basis of either the slope or the initial index of a linear equation is, however, generally inadequate, because the relative growth of an organism is not necessarily linear. The simple ratio between two linear measurements is actually variable as the organism grows. The histograms of isometric variation would be apt to be flat-topped (platykurtic) and skewed in comparison with the ideal normal distribution, if the sample were composed of individuals of different ages (text-fig. 3).

There are several methods for obtaining a regressed line of best-fit. As discussed by Teissier (1948), Kermack and Haldane (1950), Imbrie (1956), Miller and Kahn (1962), and others, the method of reduced major axis seems to be more reasonable and advantageous in biometrical studies than the conventional regression analysis y on x or x on y . In a statistical study of Cenozoic *Argopecten*, Waller (1969) found empirically that the reduced major axis fits a greater variety of point distributions than Bartlett's line.

The slope of the reduced major axis is given as:

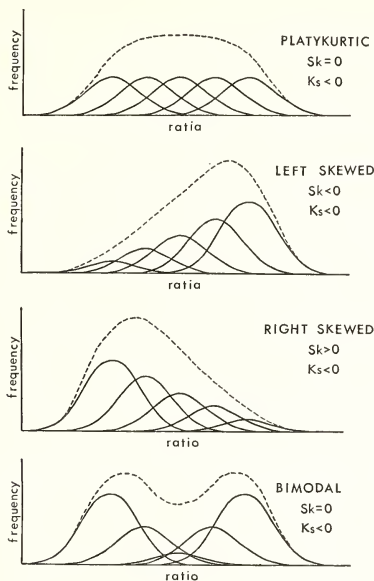
$$a = \frac{s_y}{s_x} = \sqrt{\left(\frac{\sum y^2 - \bar{y} \sum y}{\sum x^2 - \bar{x} \sum x} \right)} \quad (5)$$

where s_x and s_y are the standard deviations of x and y respectively. If the two variables of the equation (3) are substituted by the mean values of x and y , the initial index of this axis is readily determined. Incidentally, the slopes of ordinary regressed lines of y on x and x on y are given by $r \cdot s_y / s_x$ and $s_y / r \cdot s_x$ respectively, where r is the correlation coefficient between two variables. Because r is positive and smaller than 1, the gradient of the reduced major axis is equal to the geometric mean of the slopes of two ordinary regressed lines. Therefore, the reduced major axis is also intuitively more reasonable than ordinary regressed lines, if the two variables are equal in dimensions.

The initial index of the linear equation, however, may be biologically meaningless, even if it is obtained by the method of reduced major axis. Moreover, it is a great difficulty for the methods of 'initial index variation' that the state of the frequency distribution is strongly influenced by the heterogeneity of the sample, especially by the age distribution within a population. This may be a serious objection in comparing fossil populations, because the age distribution must be controlled not only by the biological condition but also by sorting in the process of sedimentation.

On the other hand, it is much more reasonable to consider that a pair of variables,

which are closely related to growth, increase approximately in accordance with a non-linear function. This relation has long been called allometry (or heterogony) (Huxley 1932, Huxley and Teissier 1936, Gould 1966). In this respect isometry is better



TEXT-FIG. 3. Hypothetical frequency distributions of heterogeneous samples, showing artificially platykurtic, skewed and bimodal tendencies.

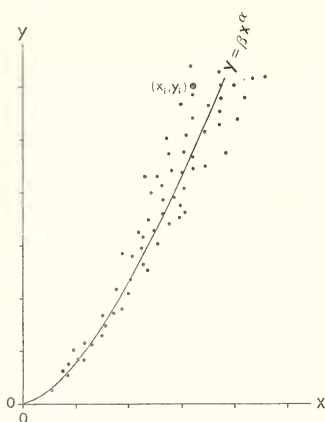
regarded as a special case of allometry. Since the pioneer work of Nomura (1926), a power equation:

$$y = \beta x^\alpha \quad (6)$$

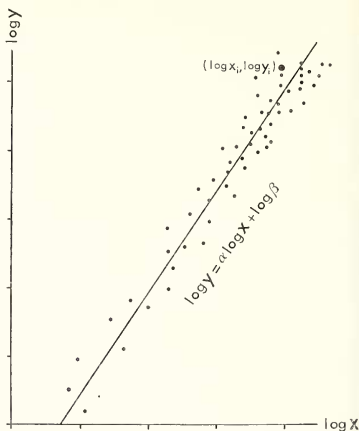
has been generally applied in the study of allometry, where the exponent α is called the specific growth ratio (or relative growth coefficient) and β the initial growth index. The adequacy of this function has been recognized empirically in many organisms.

As pointed out by some authors (e.g. Shimizu 1959, Röhrs 1961, Gould 1966), the method of study of ontogenetic allometry may be classified broadly into two categories: one is the direct examination of individual relative growth, and the other is the estimation of average relative growth deduced from many individuals of various growth stages. The former is no doubt more advantageous than the latter in many respects (Gould 1966, Obata 1967), but cannot be pursued in the study of fossils, unless the

morphology of every growth stage is preserved in an individual. Furthermore, we must treat the sample representing a population instead of the individual in a quantitative study of variation. In other words we should focus on the population allometry, paying attention also to the individual growth.



TEXT-FIG. 4. Hypothetical scatter diagram showing the concept of allometric variation of a bivariate character. The power function $y = \beta x^\alpha$ represents the average allometry for a sample.



TEXT-FIG. 5. Data of text-fig. 4 plotted as $\log x$ $\log y$ on a double logarithmic paper ($X = \log x$, $Y = \log y$).

The allometric equation (6) is just equivalent to the following linear equation:

$$\log y = \alpha \log x + \log \beta. \quad (7)$$

The equation of average ontogenetic allometry for a sample can be obtained by the regression of $\log y$ on $\log x$ (or vice versa) or more desirably by the method of reduced major axis with transformation of the original linear measurements into logarithms. Imbrie (1956) showed empirically that better results would be obtained by the method of reduced major axis and by the transformation of original bivariate data into logarithms. As exemplified in text-figs. 4 and 5, it is reasonable to consider that the deviation for each individual is well represented by some standardized distance from the line of average allometry (= reduced major axis). Although there are several different methods to measure the deviation, we like to discuss the frequency distribution of the distance collectively in terms of *allometric variation*.

REPRESENTATION OF ALLOMETRIC VARIATION

In order to recognize allometric growth it is convenient to plot the original bivariate data on a double logarithmic paper (see text-fig. 5). If a linear relation was expected

between $\log x$ and $\log y$ in the scatter diagram by intuition or high value of correlation coefficient, a power equation for the average ontogenetic allometry could be justified.

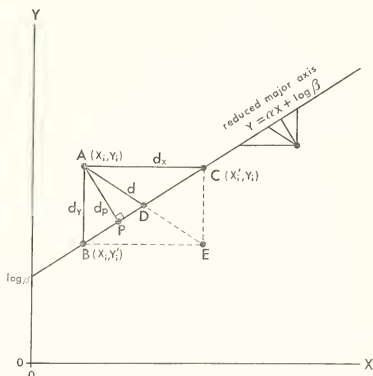
Sometimes, however, the average relative growth of an organism might better be regarded as a more complicated function. Some authors have expressed such a relation by the composition of connected lines of different slopes, where 'di or poly-phasic allometry' might be suggested. In such a case the sample may better be split, though artificially, into several classes of different ontogenetic stages in accordance with the expected 'critical point(s)', before the equation of average ontogenetic allometry is determined. Simpson, Roe, and Lewontin (1960, p. 413) and many others introduced such a procedure. In the study of average allometry, however, this state may hardly be discriminated from the gradual change of specific growth ratio. Therefore, precise examination on the individual relative growth may be also necessary for the recognition of 'poly-phasic allometry'. This problem was discussed on a theoretical and biological basis by Gould (1966). He is of the opinion that the so-called 'poly-phasic allometry' and 'critical point(s)' are usually nothing but pure artifacts of an improper procedure [personal communication from Dr. S. J. Gould (17 September 1969)].

Many authors have assumed that α is constant in one population or even in one species and that β is related principally to the individual or geographic variation. Actually, however, the value of α must vary to some extent among individuals of the same population. This is, in fact, a very troublesome problem, because the variability of α could not be clarified without a detailed analysis of individual growth. Therefore, it is not quite adequate to discuss the variation by means of the initial growth indices for individuals on the assumption of strictly parallel lines to the line of average allometry.

More adequate and easily measurable index is the actual distance from each point to the line of average allometry (the reduced major axis) on a double logarithmic scatter diagram (text-fig. 6). At least five kinds of measurements are available for the distance, namely, Y-distance, X-distance, diagonal distance, perpendicular distance, and triangle-root distance. They can be computed in the following manner:

1. Y-distance (d_Y). The distance measured parallel to the Y-axis. If $\log y$ and $\log x$ are substituted by Y and X respectively, the equation of average allometry (7) is transformed into the following simple formula:

$$Y = \alpha X + \log \beta. \quad (8)$$



TEXT-FIG. 6. Various distances from the reduced major axis. $A(X_i, Y_i)$ are the coordinates for an arbitrary individual. $d_Y (= AB)$: Y-distance; $d_X (= AC)$: X-distance; $d (= AD)$: diagonal distance; $d_P (= AP)$: perpendicular distance.

When the coordinates for the point B are designated as (X_i , Y'_i), the Y -distance for an arbitrary individual is obviously given as:

$$d_Y = Y_i - Y'_i = Y_i - \alpha X_i - \log \beta. \quad (9)$$

This index was applied already by Richards and Kavanaugh (1945) and quoted by Simpson, Roe, and Lewontin (1960) in relation to the analysis of 'di-phasic' allometry.

2. X -distance (d_X), the distance measured parallel to the X -axis. If the coordinates for the point C are defined as (X'_i , Y_i), the X -distance is represented as:

$$d_X = X'_i - X_i = \frac{Y_i - \log \beta}{\alpha} - X_i. \quad (10)$$

3. Diagonal distance (d). Imbrie (1956) suggested that one of the most reasonable indices for the deviation is the diagonal distance (\overline{AD}) from each point to the line of average allometry which is determined by the method of reduced major axis. This index could be calculated in the following manner:

$$\overline{AD} = \frac{1}{2} \overline{AE} = \frac{\sqrt{\{(\overline{AB})^2 + (\overline{AC})^2\}}}{2}.$$

Therefore,
$$d = \frac{\sqrt{(d_Y^2 + d_X^2)}}{2} = \frac{(Y_i - \alpha X_i - \log \beta) \sqrt{(\alpha^2 + 1)}}{2\alpha}. \quad (11)$$

4. Perpendicular distance (d_P). This is of course the shortest distance (\overline{AP}), which is measured perpendicularly from each point to the line of average allometry. It is actually similar to the deviation in the method of 'major axis', where the line of best-fit is obtained by minimizing the sum of the squares of the perpendicular distance. Kermack and Haldane (1950) and Imbrie (1956) pointed out its inappropriateness for the determination of the line of average allometry, because its slope may change with the unit of measurement. In discussing the deviation from the already determined line of average allometry, however, this distance seems to be one of the meaningful indices. Because ΔPBA is a similar figure to ΔABC (see text-fig. 6), the value of perpendicular distance can be determined in the following manner:

$$\frac{\overline{AP}}{\overline{AB}} = \frac{\overline{AC}}{\overline{BC}}, \quad \overline{AP} = \frac{\overline{AB} \cdot \overline{AC}}{\overline{BC}} = \frac{\overline{AB} \cdot \overline{AC}}{\sqrt{\{(\overline{AB})^2 + (\overline{AC})^2\}}}.$$

Therefore,
$$d_P = \frac{d_Y \cdot d_X}{\sqrt{(d_Y^2 + d_X^2)}} = \frac{Y_i - \alpha X_i - \log \beta}{\sqrt{(\alpha^2 + 1)}}. \quad (12)$$

5. Triangle-root distance (d_R). A reduced major axis is obtained by minimizing the sum of the areas of triangles ΔABC . Therefore, the square root of the area of each triangle may be also a useful index for the deviation, although it can be hardly measured directly on a scatter diagram

$$d_R = \sqrt{\left(\frac{\overline{AB} \cdot \overline{AC}}{2}\right)} = \sqrt{\left(\frac{d_Y \cdot d_X}{2}\right)} = \frac{Y_i - \alpha X_i - \log \beta}{\sqrt{(2\alpha)}}. \quad (13)$$

The actual values of these distances are always positive for the points above the line of average allometry and always negative for those below the line. This may be a convenient nature for further statistical treatment.

As readily recognized from the formulae (9)–(13), the values of these distances are completely proportional to one another.

$$d_Y:d_X:d_P:d_R = 1:\frac{1}{\alpha}:\frac{\sqrt{(\alpha^2+1)}}{2\alpha}:\frac{1}{\sqrt{(\alpha^2+1)}}:\frac{1}{\sqrt{2\alpha}} \quad (14)$$

Therefore, the following relation is also recognized as to the absolute values of distances for an arbitrary individual:

$$|d| \geq |d_R| \geq |d_P|. \quad (15)$$

The value of sample standard deviation also depends upon the above proportional expression (14). Consequently one would obtain essentially the same pattern of frequency distribution, whatever distances might be applied as the index. Every distance may be applicable for the representation of variation, but, we think, the diagonal distance is the best, when the two variables are the same in dimensions. However, provided that one variable is dependent on the other (whorl height of coiling shell versus volution number for example), it may be desirable to apply the Y -distance from the ordinary regressed line (Y on X).

Generally speaking, the degree of variability should be expressed on the basis of the value of sample standard deviation (s) in comparison with the sample mean (M), as Pearson's coefficient of variation (V) is defined as:

$$V = \frac{100s}{M}. \quad (16)$$

In discussing the variability of the distance from the reduced major axis, Teissier (1948) gave the following expression for the variance of diagonal distance, using the vector sum of d_Y and d_X :

$$s_d^2 = \frac{1}{2}((1-r)(s_X^2 + s_Y^2)) \quad (17)$$

where s_X and s_Y are the sample standard deviations of X and Y , and r is the correlation coefficient between the two variables. As discussed by Imbrie (1956), the value of the variability of d should be represented by the value of s_d in comparison with the joint mean of \bar{X} and \bar{Y} .

$$V_d = \frac{100s_d}{\sqrt{(\bar{X}^2 + \bar{Y}^2)}} = 100 \sqrt{\frac{(1-r)(s_X^2 + s_Y^2)}{2(\bar{X}^2 + \bar{Y}^2)}}. \quad (18)$$

(Imbrie (1956) actually gave the coefficient of variability for twice the diagonal distance of the present usage.) In applying this parameter to this coefficient we could discuss the variability of a bivariate character. It is generally supposed that the variability of diagonal distance is smaller than that of the simple ratio, if the relative growth of an organism is better represented by a power function than a linear equation.

AN EMPIRICAL EVALUATION OF ALLOMETRIC VARIATION ON THE BASIS OF ACTUAL DATA

As discussed above, it is theoretically certain that the variation of distance from the line of average allometry (reduced major axis in double logarithmic field) is a more desirable representation for a bivariate character than that of a simple ratio. We might be able to recognize the advantage of the former method also empirically, if these two methods were applied independently to the same sample. In the present article we intend to take a fossil sample of *Glycymeris* (Bivalvia) as an illustrative example.

The following is the basic information about the sample here analysed.

Specific name. *Glycymeris rotunda* (Dunker).

Locality. Loc. 4, Ugari, Fukuroi City, Shizuoka Prefecture, central Japan (collected by A. Matsukuma). This is probably the same as Loc. 525 described by Makiyama (1941).

Horizon. Hosoya tuffaceous siltstone member of the Kakegawa group.

Age. Kechienjian stage (Pliocene).

Statistical sample. Only right valves were used. The sample is composed of 203 individuals which were collected randomly from a fossil bed. They consist of individuals of quite various size, although smaller ones are relatively abundant (see text-fig. 8).

For the sake of convenience the following characterization is applied for the linear measurements and related auxiliary variables (see also text-fig. 7).

x : The length of the dental plate measured from the umbo to the posterior extremity.

y : The length of the dental plate measured from the umbo to the anterior extremity.

z : The simple ratio between the two linear measurements (y/x).

X : The common logarithms of x [$X = \log x$].

Y : The common logarithms of y [$Y = \log y$].

The values of x and y were measured strictly along the direction parallel to the basal line of ligament area. The mensuration was carried out by means of a specially designed comparator which was recently introduced by Shuto (1969, p. 49).

1. Analysis of isometric variation

From the bivariate data of the 203 individuals the following fundamental values were calculated in relation to the isometric variation:

$$\bar{z} = \frac{1}{203} \sum \frac{y_i}{x_i} = 1.16651,$$

$$s_z = \sqrt{\left\{ \frac{\sum (z_i - \bar{z})^2}{202} \right\}} = 0.11788,$$

$$V_z = \frac{100s_z}{\bar{z}} = 10.105.$$

As shown in text-fig. 9 and also in Table 1, the null hypothesis for goodness-of-fit of the actual data to a normal distribution is accepted.

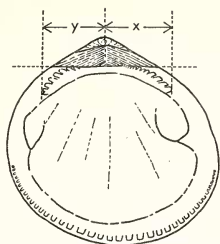
Kurtosis is the property of being more pointed or flatter than a normal curve. As noted by Simpson, Roe, and Lewontin (1960, pp. 146-7), the best measure of kurtosis (K_s) is given as follows:

$$K_s = \frac{\sum (x_i - \bar{x})^4}{Ns^4} - 3, \quad (19)$$

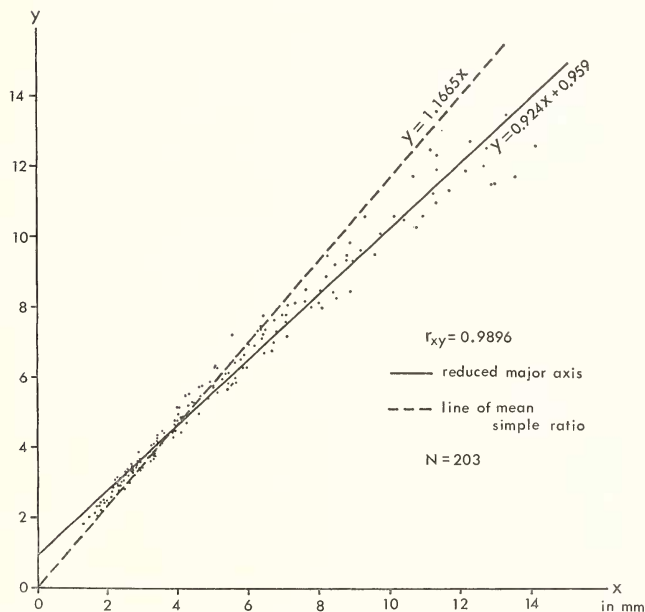
where N is the number of individuals and s the sample standard deviation. In this case we obtained the following value as the kurtosis coefficient for the distribution of z :

$$K_s = \frac{\sum (z_i - \bar{z})^4}{203s_z^4} - 3 = -0.5257.$$

Because the value is negative, the distribution is flatter than the normal curve and may be said to be platykurtic. We presume that this tendency is partly due to the gradual



TEXT-FIG. 7. Linear measurements (in millimetres) adopted in the present study.
 x : The length of dental plate from the umbo to the posterior extremity; y : The length of dental plate umbo to the anterior extremity.



TEXT-FIG. 8. Scatter diagram for a sample of *Glycymeris rotunda* (Dunker) from a fossil bed of the Kakegawa group at Loc. 4, Ugari, Fukuroi City, Shizuoka Pref., Japan. The reduced major axis does not fit to the original bivariate data especially in small individuals.

change of z through the growth of each individual. In fact, the average value of z is much larger in smaller individuals than in larger ones.

Skewness is a parameter showing the property of asymmetric frequency distribution. It is commonly given as:

$$Sk = \frac{\sum (x_i - \bar{x})^3}{N s^3}. \quad (20)$$

TABLE 1

Calculation of χ^2 for goodness-of-fit of the observations of z to a normal distribution

	Simple ratio (y/x)	Normal probability	Observed frequency (O_i)	Expected frequency (E_i)	$\frac{(O_i - E_i)^2}{E_i}$
$\bar{z} - 3s_z \sim \bar{z} - 2s_z$	(0.8128~0.9307)	0.0215	5	4.36	0.0939
$\bar{z} - 2s_z \sim \bar{z} - s_z$	(0.9307~1.0486)	0.1359	32	27.59	0.7049
$\bar{z} - s_z \sim \bar{z}$	(1.0486~1.1665)	0.3413	64	69.28	0.4024
$\bar{z} \sim \bar{z} + s_z$	(1.1665~1.2844)	0.3413	71	69.28	0.0427
$\bar{z} + s_z \sim \bar{z} + 2s_z$	(1.2844~1.4023)	0.1359	30	27.59	0.2105
$\bar{z} + 2s_z \sim \bar{z} + 3s_z$	(1.4023~1.5202)	0.0215	1	4.36	2.5894
Total		0.9974	203	202.46	4.0438

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 4.04 \quad [\text{with 3 degrees of freedom}].$$

$$\chi^2_{0.05}(v=3) = 7.81.$$

$$0.25 < P < 0.30.$$

For reference: If the data of z are regrouped to interval $0.5s_z$, the result of χ^2 test is as follows:

$$\chi^2 = 11.95 \quad [\text{with 7 degrees of freedom}]$$

$$\chi^2_{0.05}(v=7) = 14.07.$$

$$0.10 < P < 0.15.$$

This coefficient is positive in a right-skewed distribution and negative in a left-skewed distribution. In the present sample the following value was obtained for the skewness of the distribution of z .

$$Sk = \frac{\sum (z_i - \bar{z})^3}{203 s_z^3} = -0.2250.$$

The left-skewed distribution is possibly related to the sample heterogeneity especially the relative abundance of small individuals.

Incidentally, we obtained the following values as to the correlation coefficient between the two linear measurements and the reduced major axis.

$$r_{xy} = 0.9896 \quad [0.9862-0.9921 \text{ for the } 95 \text{ per cent confidence interval}],$$

$$y = 0.924x + 0.959.$$

2. Analysis of allometric variation

The reduced major axis (best-fitted line of average allometry) for the present sample was determined in the following manner:

$$\bar{X} = 0.64791, \quad \bar{Y} = 0.71253,$$

$$s_X = 0.25213, \quad s_Y = 0.21411.$$

The following value was obtained for the correlation coefficient between X and Y :

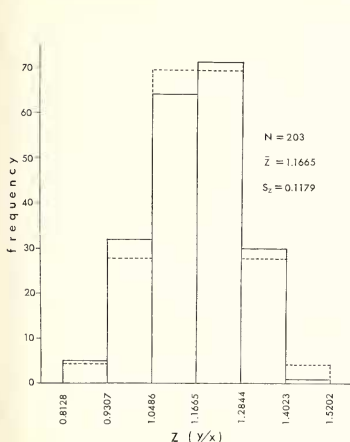
$$r_{XY} = 0.9948 [0.9931-0.9961 \text{ for the } 95\% \text{ confidence interval}].$$

Therefore, the correlation between the two variables is very significant. The slope of reduced major axis is given by the formula (5),

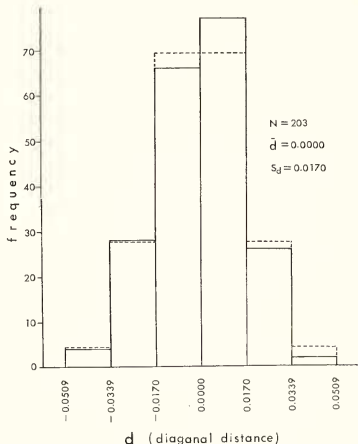
$$\alpha = \frac{s_Y}{s_X} = \frac{0.21411}{0.25213} = 0.84920.$$

From the formula (8),

$$\log \beta = Y - \alpha X.$$



TEXT-FIG. 9. Frequency distribution of $z (= y/x)$. The broken lines show the expected values in theoretical normal distribution.



TEXT-FIG. 10. Frequency distribution of d (diagonal distance from the reduced major axis). The broken lines show the expected values in theoretical normal distribution.

By the substitution of \bar{X} and \bar{Y} ,

$$\log \beta = 0.71253 - 0.84920 \times 0.64791 = 0.16232.$$

Therefore, the reduced major axis for the present sample is represented by the following equation:

$$Y = 0.84920X + 0.16232 \quad \text{or} \quad y = 1.4532x^{0.84920}.$$

The standard error of the slope (α) is given as:

$$\sigma_\alpha = \alpha \sqrt{\left(\frac{1-r^2}{N}\right)} = 0.84920 \times \sqrt{\left(\frac{1-0.9948^2}{203}\right)} = 0.00608.$$

Consequently, the interval of α with 95% confidence is:

$$\alpha \pm 1.96\sigma_\alpha = 0.84920 \pm 0.01192.$$

Therefore, the relative growth is certainly allometric. The diagonal distance (d), which

is measured from each point to this reduced major axis in a double logarithmic field, is given as:

$$d = \frac{(Y_i - \alpha X_i - \log \beta) \sqrt{(\alpha^2 + 1)}}{2\alpha} = 0.77245(Y_i - 0.84920X_i - 0.16232).$$

The computation of d for 203 individuals results in:

$$\bar{d} = 0.00002 \quad [\text{negligible}]$$

$$s_d = 0.01697.$$

TABLE 2

Calculation of χ^2 for goodness-of-fit of the observations of d to a normal distribution

Diagonal distance (d)	Normal probability	Observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$ E_i
$\bar{d} - 3s_d \sim \bar{d} - 2s_d$ ($-0.0509 \sim -0.0339$)	0.0215	4	4.36	0.0297
$\bar{d} - 2s_d \sim \bar{d} - s_d$ ($-0.0339 \sim -0.0170$)	0.1359	28	27.59	0.0061
$\bar{d} - s_d \sim \bar{d}$ ($-0.0170 \sim 0$)	0.3413	66	69.28	0.1553
$\bar{d} \sim \bar{d} + s_d$ ($0 \sim 0.0170$)	0.3413	77	69.28	0.8603
$\bar{d} + s_d \sim \bar{d} + 2s_d$ ($0.0170 \sim 0.0339$)	0.1359	26	27.59	0.0916
$\bar{d} + 2s_d \sim \bar{d} + 3s_d$ ($0.0339 \sim 0.0509$)	0.0215	2	4.36	1.2774
Total	0.9974	203	202.46	2.4204

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 2.42 \quad [\text{with 3 degrees of freedom}].$$

$$\chi^2_{0.05(v=3)} = 7.81.$$

$$0.40 < P < 0.50.$$

For reference: If the data of d are regrouped to interval $0.5s_d$, the result of χ^2 tests is as follows:

$$\chi^2 = 10.00 \quad [\text{with 7 degrees of freedom}].$$

$$\chi^2_{0.05(v=7)} = 14.07.$$

$$0.15 < P < 0.20.$$

If the formula (18) were applied for the estimation of variability, the coefficient of variation might be shown:

$$V_d = \frac{100s_d}{\sqrt{(\bar{X}^2 + \bar{Y}^2)}} = \frac{1.697}{\sqrt{(0.6479^2 + 0.7125^2)}} = 1.762.$$

As shown in text-fig. 10 and also Table 2, the frequency distribution of d is also regarded as normal. The distribution seems to be also platykurtic in view of the negative value of kurtosis coefficient:

$$Ks = \frac{\sum(d_i - \bar{d})^4}{203s_d^4} - 3 = -0.3570.$$

The skewness for the distribution of d is given as:

$$Sk = \frac{\sum(d_i - \bar{d})^3}{203s_d^3} = -0.0292.$$

It means a slightly left-skewed distribution.

3. Comparisons and discussions

Now, we intend to compare the result obtained by the method of diagonal distance with that by the method of simple ratio. The obtained value of sample correlation coefficient between X and Y is significantly higher than that between x and y . We presume that the appropriateness of the index for a bivariate character is indicated to a certain extent by the smallness of the coefficient of variation. Although direct comparison may not be meaningful, the value of V_d is obviously much smaller than V_z .

The result of χ^2 test for the frequency distribution of d supports the null hypothesis of a normal distribution. The value of χ^2 (with the same degrees of freedom) is more or less smaller and the probability for a normal distribution higher than those for the distribution of z (Tables 1 and 2). In this respect the pattern of histogram of d is more safely assumed to fit a theoretical normal curve than is that of z .

As to the coefficients of kurtosis and skewness, a similar assumption is possible. The values of Ks and Sk indicate that the frequency distribution of d is less platykurtic and less left-skewed than that of z .

As indicated in text-fig. 8, neither the isometric line, $y = 1.16651x$, nor the reduced major axis, $y = 0.924x + 0.959$, fits the original bivariate data. On the other hand, as shown in text-fig. 11, the reduced major axis in double logarithmic field seems to represent well the relative growth of this sample. Therefore, so far as the present sample is concerned, monophasic allometry is well recognized. Judging from the comparative data aforementioned, it is concluded empirically that the diagonal distance from the reduced major axis is a more desirable index for the expression of a bivariate character than is the simple ratio between two variables.

STATISTICAL DEFINITION OF 'ISOMETRY' AND 'ALLOMETRY'

In order to identify or to discriminate samples purely statistically, Imbrie (1956) and Miller and Kahn (1962) introduced a method of significance test for the difference of two reduced major axes. It is, in fact, a very useful method, if the relative growth of each sample shows monophasic allometry.

The standard errors of the slopes or reduced major axes are given as:

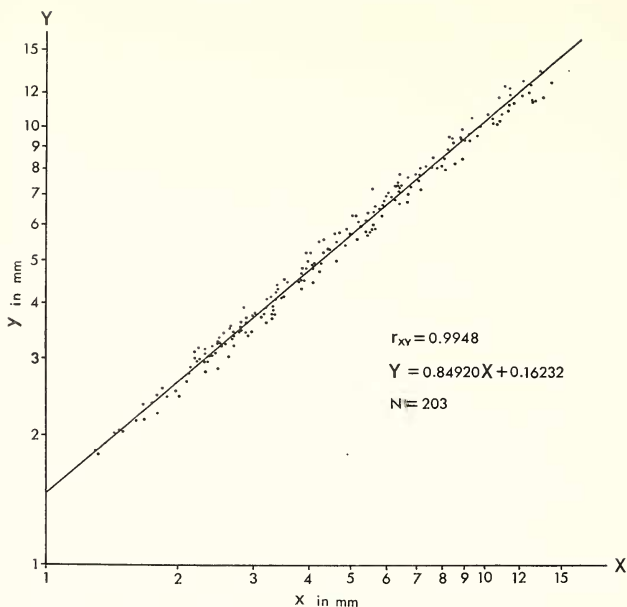
$$\sigma_{\alpha_1} = \alpha_1 \sqrt{\left(\frac{1-r_1^2}{N_1} \right)} \quad (21)$$

$$\sigma_{\alpha_2} = \alpha_2 \sqrt{\left(\frac{1-r_2^2}{N_2} \right)} \quad (22)$$

where α_1 and α_2 are the slopes of two reduced major axes, r_1 and r_2 are the correlation coefficients between two variables and N_1 and N_2 are the numbers of individuals in respective samples. Provided that the numbers of individuals are not too small, one can recognize by means of the following value whether the difference of slopes is significant or not

$$Z = \frac{\alpha_1 - \alpha_2}{\sqrt{(\sigma_{\alpha_1}^2 + \sigma_{\alpha_2}^2)}} \quad (23)$$

If $|Z| < 1.96$, the difference of slopes is not significant with 95% confidence, and if $|Z| > 1.96$, the two samples can be discriminated by the difference of specific growth ratio.



TEXT-FIG. 11. Double logarithmic scatter diagram for a sample of *Glycymeris rotunda* (Dunker) [the same sample as shown in text-fig. 8.] Because the reduced major axis fit well to the data, the relative growth can be regarded as monophasic.

Applying the method of this significance test, we can define statistically the terms of 'isometry' and 'allometry'. In the case of ideal isometry, though such a state is actually non-existent, the slope is, of course, strictly equal to 1 without error. Consequently, if the value of Z is calculated by the following equation:

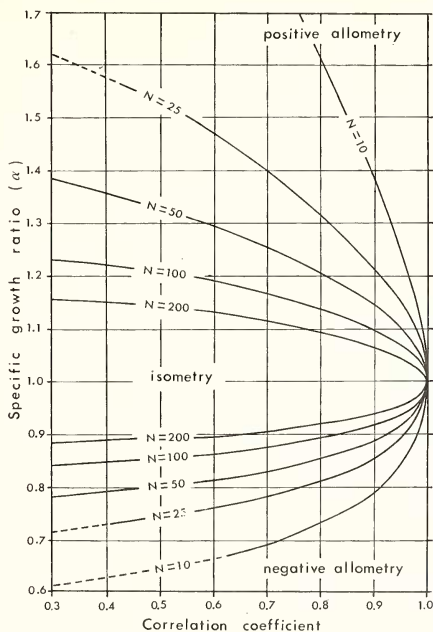
$$Z = \frac{\alpha - 1}{\sigma_{\alpha}} = \frac{\alpha - 1}{\alpha \sqrt{\{(1 - r^2)/N\}}} \quad (24)$$

one can judge with 95% confidence whether the relative growth is isometric or allometric, as follows:

- if $-1.96 < Z < 1.96$, the null hypothesis for 'isometry' would be accepted,
- if $Z < -1.96$, negative allometry would be suggested, and
- if $Z > 1.96$, positive allometry would be suggested.

Because the simple ratio between two variables is an index on the assumption of isometry, the allometric variation is certainly a more reasonable representation than the

isometric one, if the null hypothesis were rejected by the above-mentioned significance test.



TEXT-FIG. 12. Statistical discrimination of 'isometry' and 'allometry'. N : number of individuals. Confidence 95%.

CONCLUDING REMARKS

It would be logically inadequate to discuss the variation of a bivariate character without considering the relative growth of the organism, if the sample is composed of individuals of various growth stages. If the simple ratio between two linear measurements were used as the index, we might have much trouble in analysing the real state of variation, especially in applying advanced statistical techniques for the recognition of a normal distribution of character and random sampling as well as taxonomic identification or discrimination. As pointed out by Simpson, Roe, and Lewontin (1960), a platykurtic and skewed distribution may be due to the heterogeneity of sample. In the analysis of bivariate characters it is necessary to apply an index which is less influenced by the age distribution.

On the other hand, if the organism proved to grow approximately in accordance with a simple power function, the individual variation within a sample could be analysed more reasonably by means of the distance from each point to the line of average allometry on a double logarithmic scatter diagram. Although at least five kinds of distance (d_Y , d_X , d , d_P , and d_R in this paper) are available for the index, this representation may be collectively called allometric variation. It is believed that this representation is more advantageous than the conventional method of simple ratio in the following respects:

1. The coefficient of variation may be reasonably small in comparison with that of simple ratio.
2. The standard deviation, skewness, and kurtosis as well as the shape of the histogram may be scarcely, if not at all influenced by the age distribution in a population and other sample heterogeneity.
3. Normal frequency distribution of bivariate characters would be recognized more reasonably on a firm basis.
4. The tendency for artificially skewed and platykurtic distribution could be avoided.
5. A comparison between different samples could be well based on the significance test for the difference of slopes and positions of reduced major axes. This procedure was fully explained by Imbrie (1956, pp. 235-8). The contingency of age distribution in samples might be negligible also in this case.
6. As pointed out by Gould (1966), the study of average allometry may not be an adequate approach to individual ontogeny, when strong natural selection takes place. The present analysis of allometric variation, however, is considered to be sometimes informative also for the consideration of the influence of natural selection and environmental factors.

The advantage of this method is also recognized empirically by a comparative study in which a sample of fossil *Glycymeris* from the Pliocene of Japan is treated as an illustrative example.

As noted by Kotaka (1953), the method of rejection ellipse may be a useful method for taxonomic identification and discrimination of fossil populations. In discussing the relation of characters on the basis of simple ratios, however, this approach may bear some difficulty, unless the ontogenetic transformation and the age distribution of samples are sufficiently considered.

On the contrary it may be an objection that the present method is somewhat time-consuming to justify its general application. More complicated techniques would be required in the organisms showing 'polyphasic' allometry. The recent rapid development of computers, however, would make the application easy, and, we believe, biometricians should not mind taking this trouble.

Acknowledgements. We express our sincere thanks to Dr. Stephen Jay Gould of Harvard University and Dr. Norman D. Newell of the American Museum of Natural History for their invaluable suggestion about the method of the reduced major axis and for their kind supervision of the manuscript. Acknowledgements are also due to Professor Tatsuro Matsumoto, Professor Ryuzo Toriyama, Dr. Tsugio Shuto, Mr. George Kato, and Mr. Tomowo Ozawa of the Department of Geology, Kyushu University; Professor Akio Kudo, Miss Kami Honda, and Mr. Kiyoshi Kawazu of the Department of Mathematics of the same university; and Dr. Ikuwo Obata of the National Science Museum (Tokyo) for their kind assistance and fruitful discussions.

REFERENCES

- GOULD, S. J. 1966. Allometry and size in ontogeny and phylogeny. *Biol. Rev.* **41**, 587-640.
- HUXLEY, J. S. 1932. *Problems of relative growth*. 276 pp. London.
- and TEISSIER, G. 1936. Terminology of relative growth. *Nature*, **137**, 780-1.
- IMBRIE, J. 1956. Biometrical methods in the study of invertebrate fossils. *Bull. Am. Mus. Nat. Hist.* **108**, 215-52.
- KERMACK, K. A. and HALDANE, J. B. S. 1950. Organic correlation and allometry. *Biometrika*, **37**, 30-41.
- KOTAKA, T. 1953. Variation of Japanese *Anadara granosa*. *Trans. Proc. Pal. Soc. Japan* [N.S.], **10**, 31-6.
- MAKIYAMA, J. 1941. The Kechienjian fauna series no. 1: Evolution of gastropod genus *Siphonalia* etc. *Mem. Fac. Sci., Kyoto Imp. Univ.* [B], **16**, 75-93.
- MAYR, E. 1969. *Principles of systematic zoology*. 428 pp. New York.
- LINSLEY, E. G., and USINGER, R. L. 1953. *Methods and principles of systematic zoology*. 336 pp. New York.
- MILLER, R. L. and KAHN, J. S. 1962. *Statistical analysis in the geological sciences*. 483 pp. New York.
- NOMURA, E. 1926. An application of $a = kb^x$ in expressing the growth relation in the freshwater bivalve *Sphaerium heterodon* Pils. *Sci. Rep. Tohoku Imp. Univ.* [4], **2**, 57-62.
- OBATA, I. 1967. Notes on relative growth in palaeontology. *Fossils* (14), 20-39 [in Japanese].
- OMORI, M. and UTASHIRO, T. 1954. Some problems on the variation of *Pecten albicans*. *Seibutsu-kagaku*, special volume (for evolution), 16-22 [in Japanese].
- RICHARDS, O. W. and KAVANAUGH, A. J. 1945. The analysis of growing form. pp. 180-230. In *Essays on growth and form*, LE GROS CLARK, W. E. and MEDAWAR, P. B., eds., Oxford.
- RÖHRS, M. 1961. Allometrieforschung und biologische Formanalyse. *Zeit. Morph. Anthropol.* **51**, 289-321.
- SHIMIZU, S. 1959. *Relative growth*. 287 pp. Tokyo [in Japanese].
- SHUTO, T. 1969. Neogene gastropods from Panay Island, the Philippines. *Mem. Fac. Sci., Kyushu Univ.* [D], **19**, (1), 1-250, pl. 1-24.
- SIMPSON, G. G., ROE, A., and LEWONTIN, R. C. 1960. *Quantitative zoology*, revised ed. 440 pp. New York.
- TEISSIER, G. 1948. La relation d'allométrie: sa signification statistique et biologique. *Biometrics*, **4**, 14-48.
- WALLER, T. R. 1969. The evolution of the *Argopecten gibbus* stock (Mollusca: Bivalvia), with emphasis on the Tertiary and Quaternary species of eastern North America. *Paleont. Soc. Mem.* **2**, [J. Paleont., **43**, Suppl. to No. 5], 1-125, pl. 1-8.

ITARU HAYAMI
AKIHIKO MATSUKUMA
Department of Geology
Faculty of Science
Kyushu University
Fukuoka 812, Japan

Typescript received 12 February 1970